PLOS ONE

RESEARCH ARTICLE

Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review

Olalekan Lee Aiyegbusi[1,2], Derek Kyte[1,2]*, Paul Cockwell[1,3], Tom Marshall[1,2], Adrian Gheorghe[4], Thomas Keeley[5], Anita Slade[1,2], Melanie Calvert[1,2]

1 Centre for Patient Reported Outcomes Research, University of Birmingham, Edgbaston, Birmingham, United Kingdom, **2** Institute of Applied Health Research, University of Birmingham, Edgbaston, Birmingham, United Kingdom, **3** Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Queen Elizabeth Hospital Birmingham, Edgbaston, Birmingham, United Kingdom, **4** Oxford Policy Management Ltd, Oxford, United Kingdom, **5** PAREXEL International, London, United Kingdom

* d.g.kyte@bham.ac.uk

Check for updates

OPEN ACCESS

Citation: Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Gheorghe A, Keeley T, et al. (2017) Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: A systematic review. PLoS ONE 12(6): e0179733. https://doi.org/10.1371/journal.pone.0179733

Editor: Kathrin Eller, Medizinische Universitat Graz, AUSTRIA

Received: April 5, 2017

Accepted: June 2, 2017

Published: June 21, 2017

Copyright: © 2017 Aiyegbusi et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This project was funded as part of the Health Foundation's Improvement Science Programme. Website: http://www.health.org.uk/ OLA is the recipient of the Health Foundation's PhD Awards for Improvement Science (Ref: 7452). The Health Foundation is an independent charity committed to bringing about better health and

Abstract

Background

Patient-reported outcome measures (PROMs) can provide valuable information which may assist with the care of patients with chronic kidney disease (CKD). However, given the large number of measures available, it is unclear which PROMs are suitable for use in research or clinical practice. To address this we comprehensively evaluated studies that assessed the measurement properties of PROMs in adults with CKD.

Methods

Four databases were searched; reference list and citation searching of included studies was also conducted. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist was used to appraise the methodological quality of the included studies and to inform a best evidence synthesis for each PROM.

Results

The search strategy retrieved 3,702 titles/abstracts. After 288 duplicates were removed, 3,414 abstracts were screened and 71 full-text articles were retrieved for further review. Of these, 24 full-text articles were excluded as they did not meet the eligibility criteria. Following reference list and citation searching, 19 articles were retrieved bringing the total number of papers included in the final analysis to 66. There was strong evidence supporting internal consistency and moderate evidence supporting construct validity for the Kidney Disease Quality of Life-36 (KDQOL-36) in pre-dialysis patients. In the dialysis population, the KDQOL-Short Form (KDQOL-SF) had strong evidence for internal consistency and structural validity and moderate evidence for test-retest reliability and construct validity while the KDQOL-36 had moderate evidence of internal consistency, test-retest reliability and construct validity. The End Stage Renal Disease-Symptom Checklist Transplantation Module

health care for people in the UK. The funder did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. AG is employed by Oxford Policy Management Ltd. Oxford Policy Management Ltd provides support in the form of salary for Dr. Gheorghe, but did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. TK is employed by PAREXEL International. PAREXEL International provides support in the form of salary for Dr. Keeley, but did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. MC receives consultancy fees from Ferring and Astellas Pharma. Ferring and Astellas Pharma did not have any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of all the authors are articulated in the 'author contributions' section.

Competing interests: OLA is funded by the Health Foundation's PhD Awards for Improvement Science. DK is funded by the National Institute for Health Research (NIHR) Fellowship. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, Health Education England or the Department of Health. The funder had no role with respect to study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication. TM is partly funded by the NIHR through the Collaborations for Leadership in Applied Health Research and Care for West Midlands (CLAHRC-WM). This paper presents independent research and the views expressed in this publication are not necessarily those of the NIHR, the Department of Health, NHS Partner Trusts, University of Birmingham or the CLAHRC-WM Management Group. AG is employed by Oxford Policy Management Ltd. TK is employed by PAREXEL International. MC reports grants from the NIHR, Macmillan Cancer Support, personal fees from Ferring and Astellas Pharma, outside the submitted work; and co-Chairs the International Society for Quality of Life Research Best Practices for PROs in Trials Taskforce. These affiliations do not alter our adherence to PLOS ONE policies on sharing data and materials.

(ESRD-SCLTM) demonstrated strong evidence for internal consistency and moderate evidence for test-retest reliability, structural and construct validity in renal transplant recipients.

Conclusions

We suggest considering the KDQOL-36 for use in pre-dialysis patients; the KDQOL-SF or KDQOL-36 for dialysis patients and the ESRD-SCLTM for use in transplant recipients. However, further research is required to evaluate the measurement error, structural validity, responsiveness and patient acceptability of PROMs used in CKD.

Introduction

Chronic kidney disease (CKD) is a global health issue [1]. It affects up to 16% of the adult population in the developed world and is associated with increased morbidity and mortality that is directly related to severity [2, 3]. CKD also has major healthcare economic costs [4, 5]. For the National Health Service (NHS) in England, the total estimated expenditure attributable to CKD between 2009 and 2010 was over a billion pounds; renal replacement therapy (RRT) accounted for approximately half of this expenditure [6]. In the US, it is estimated that managing stages 3 and 4 CKD cost the Medicare fee-for-service (FFS) approximately $44.4 billion annually [7].

Early symptoms of CKD, such as fatigue, are usually non-specific [8]. However patients with more advanced CKD often report multiple 'clusters' of symptoms including drowsiness, pain, pruritus and dry skin [9]. This overall symptom burden may have a negative impact on the perceived health-related quality of life (HRQOL) of patients with end-stage renal disease [10, 11].

HRQOL can be assessed using self-administered, validated questionnaires known as patient-reported outcome measures (PROMs) [12]. PROMs have a wide variety of applications ranging from clinical trials [13] to product labelling [14] and routine clinical care [15, 16]. There is an increasing awareness that PROMs may have a future role in the management of patients with kidney disease, including integration into routine practice; for example, through monitoring patients for symptoms or changes in HRQOL that may require an intervention [17].

If any benefit is to be derived from the use of PROMs, it is important that they are well validated to ensure that they actually measure what they are supposed to measure, produce consistent results and capture all aspects of the construct(s) under investigation that matter to the target population if any benefit is to be derived from their use [14].

A systematic review by Gibbons and Fitzpatrick [18] evaluated the measurement properties of PROMs used in the CKD population, but this was conducted over six years ago. There have been methodological advances since then [19, 20] and it is reasonable to assume that new research has been published [21]. The review was restricted to studies published in English which might have excluded potentially relevant papers [18]. In addition, the review did not report evaluating the methodological quality of the selected studies. It is vital that the methodological quality of studies evaluating the measurement properties of PROMs is assessed to ensure that conclusions about the reliability and validity of the measures are dependable [22] as these have a potential impact on clinical practice and health policy [23].

Therefore, we have evaluated the methodological quality of the selected studies using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)

checklist [22] and used the findings to inform our evidence synthesis. Thus providing the best evidence possible, to inform the selection of PROMs for monitoring the symptoms of CKD and its treatment effects in pre-dialysis, dialysis and renal transplant patients.

Methods

Design

This systematic review was conducted and reported according to a registered and published protocol (PROSPERO registration number: CRD42016035554) (See S1 Text. Review Protocol) [24] and written in compliance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines (See S2 Text. PRISMA Checklist) [25]. We also considered the findings of the review of systematic reviews by Terwee et al. [26].

Search strategy

Relevant databases including MEDLINE (Ovid), EMBASE (Ovid), PsycINFO (Ovid) and CINAHL Plus (EBSCO) were systematically searched from inception to 21st December 2015 without language restrictions [24].

The search strategy was initially developed for MEDLINE and subsequently adapted for the other databases (See S3 Text. Search Strategy). Two existing search filters [27, 28] were combined with key terms generated by the review team for renal disease and its treatment modalities. An information specialist at the Institute of Applied Health Research, University of Birmingham, was consulted during the process.

Search records were downloaded into Endnote X7 and duplicates removed. In addition, the UK Renal Registry website was searched to 17th May 2016.

Screening process

All titles and abstracts were screened independently by two reviewers (OLA and TK/AG). Full-text articles were obtained for studies potentially meeting the eligibility criteria and were independently reviewed by the same reviewers. Reasons for exclusion were documented. Hand searching of reference lists and citation searching of the included papers was also conducted. At all stages, disagreements regarding eligibility were resolved through discussion and, if necessary, consultation with a third reviewer (MC/DK).

Selection of studies

Studies were included if they: (1) focused on PROMs used specifically for measuring HRQOL and/or CKD symptoms (the constructs of interest) in any CKD population; and (2) reported either the development or evaluation of one or more psychometric properties of a PROM [24].

Articles excluded were clinical trial reports, editorials, reviews and conference abstracts. In addition, studies that focussed on clinician-assessed instruments, PROMs developed for use in patients with acute kidney injury or in patients below 18 years of age were excluded.

Data extraction

Data from selected studies were extracted independently by two reviewers (OLA and AS) using a pre-designed data collection form and cross-checked for accuracy.

The following data were extracted where available

1. Characteristics of study populations
2. Questionnaire characteristics

3. Evidence regarding measurement properties as defined by Mokkink et al.[29] namely: reliability (test-retest reliability, internal consistency, measurement error); validity (content validity, construct validity (including hypothesis testing, structural validity and cross-cultural validity); responsiveness of questionnaires to changes over time; the setting and purpose for which the questionnaires were administered and details regarding their interpretability; operational characteristics including patient acceptability and mode and feasibility of administration; and details regarding patient involvement in the PROM development or validation process.

Appraisal of the methodological quality of selected studies

Following the selection process, the quality of the included papers was assessed by two reviewers (OLA/AS) using a validated critical appraisal tool for studies of health measurement instruments: the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist [20, 22]. The COSMIN checklist is designed to evaluate the methodological quality of studies of psychometric properties [22]. It comprises of mini-checklists A to I (otherwise known as 'boxes') which correspond to each measurement property (See Table 1). Some measurement properties are named and defined differently by different authors, therefore the COSMIN definitions [29] were used to ascertain which measurement properties were evaluated by the studies. The COSMIN checklist is intended for use as a modular tool meaning that the mini-checklists (boxes) to be completed for each study will be determined by the measurement properties evaluated by the study [29]. Each mini-checklist has a set of quality items/questions which were rated individually using the COSMIN 4-point scale as 'excellent', 'good', 'fair' or 'poor'.

An item is rated 'excellent' when there is evidence that the methodological quality of the study in relation to the item is adequate [20]. An item is rated 'good' when relevant information is not reported in an article, but it can be assumed that the methodological quality is adequate [20]. An item is rated 'fair' if there is doubt about the adequacy of the study's methodological quality in relation to that item [20]. Finally, an item is rated 'poor' when there is evidence that the methodological quality of the study in relation to that particular item is inadequate [20]. For example, a small sample size was considered poor methodological quality in all the mini-checklists. A sample size $\geq$100 was considered 'excellent', 50–99 'good', 30–49 'fair', and $<$30 'poor' [20].

The 'worst score counts' method was used to determine the methodological quality of each paper per measurement property [20]. This meant taking as the overall score for each measurement property, the lowest rating given to any item within the respective mini-checklist [20].

Reviewers consulted a third author (MC/DK) if they were unable to reach a consensus at any point during the assessment.

Data synthesis

The quality criteria developed by Terwee et al. [30] was used to rate the results for each measurement property per study as either 'positive' (+), 'indeterminate' (?) or 'negative' (-). For example, structural validity was rated as positive, if the factors identified after performing a factor analysis were reported to explain at least 50% of variance. If the factors explained $<$50% of variance structural validity was rated as negative. The indeterminate rating was given if the percentage of variance explained was not reported. (See Table 1) [30].

An evidence synthesis across studies was then conducted for measurement properties reported for each PROM using another set of criteria (See Table 2) [19]. At this stage, the

Table 1. Quality criteria for measurement properties.

Property	Rating †	Quality Criteria
Reliability		
Internal consistency	+	Cronbach's alpha(s) ≥ 0.70
	?	Cronbach's alpha not determined or dimensionality unknown
	-	Cronbach's alpha(s) < 0.70
Reliability	+	ICC / weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80
	?	Neither ICC / weighted Kappa, nor Pearson's r determined
	-	ICC / weighted Kappa < 0.70 OR Pearson's r < 0.80
Measurement error	+	MIC > SDC OR MIC outside the LOA
	?	MIC not defined
	-	MIC ≤ SDC OR MIC equals or inside LOA
Validity		
Content validity	+	All items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement AND the questionnaire is considered to be comprehensive
	?	Not enough information available
	-	Not all items are considered to be relevant for the construct to be measured, for the target population, and for the purpose of the measurement OR the questionnaire is considered not to be comprehensive
Structural validity	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	-	Factors explain < 50% of the variance
Hypothesis testing	+	Correlations with instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses AND correlations with related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with instruments measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR correlations with related constructs are lower than with unrelated constructs
Cross-cultural validity	+	No differences in factor structure OR no important DIF between language versions
	?	Multiple group factor analysis not applied AND DIF not assessed
	-	Differences in factor structure OR important DIF between language versions
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥ 0.70
	?	No convincing arguments that gold standard is "gold"
	-	Correlation with gold standard < 0.70
Responsiveness		
Responsiveness	+	Correlation with changes on instruments measuring the same construct ≥ 0.50 OR at least 75% of the results are in accordance with the hypotheses OR AUC ≥ 0.70 AND correlations with changes in related constructs are higher than with unrelated constructs
	?	Solely correlations determined with unrelated constructs
	-	Correlations with changes on instruments measuring the same construct < 0.50 OR < 75% of the results are in accordance with the hypotheses OR AUC < 0.70 OR correlations with changes in related constructs are lower than with unrelated constructs

MIC = minimal important change, SDC = smallest detectable change, LoA = limits of agreement

ICC = intraclass correlation coefficient, DIF = differential item functioning, AUC = area under the curve

† + = positive rating,? = indeterminate rating

- = negative rating

(Reproduced with permission from Caroline Terwee, COSMIN)

https://doi.org/10.1371/journal.pone.0179733.t001

overall level of evidence for each PROM was provided by one or more studies, taking into account their methodological quality [20]. The overall level of evidence for each measurement property was graded as 'strong', 'moderate', 'limited' 'unknown' or 'conflicting' [19]. For example, a measurement property was graded as 'strong' if at least one study had 'excellent'

Table 2. Levels of evidence for the quality of the measurement property.[19, 30].

Level of Evidence	Rating†	Criteria
Strong	+++ or —-	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or —	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	+/-	Conflicting findings
unknown	?	Only studies of poor methodological quality

† + = positive rating

? = unknown rating

- = negative rating

(Reproduced with permission from Caroline Terwee, COSMIN)

https://doi.org/10.1371/journal.pone.0179733.t002

methodological quality or at least two studies had 'good' methodological qualities (See Table 2).

Results

The search strategy retrieved 3,702 titles/abstracts. After 288 duplicates were removed, 3,414 abstracts were screened and 71 full-text articles were retrieved for further review. Of these, 24 full-text articles were excluded for various reasons (Fig 1). Following reference list and citation searching, 19 articles were retrieved bringing the total number of papers included in the final analysis to 66. Strength of agreement between the reviewers, calculated using Cohen's Kappa Statistic [31], was good (OLA/AG = 0.889, OLA/TK = 0.863).

Description of tables

Table 1 presents the quality criteria used to evaluate the measurement properties for each PROM while the criteria used for the evidence synthesis are in Table 2. A brief description of the PROMs evaluated in this review is presented in Table 3. The methodological qualities of these studies are summarized in Table 4. Table 5 presents the synthesis of the overall level of evidence for each PROM. The results reported by the included studies were extracted and are summarized in S1 Table. Summary of study results. The characteristics of the included studies are presented in S2 Table. Characteristics of included studies. The PROMs evaluated in this review are summarized in S3 Table. Characteristics of included PROMs. S1–S3 Tables have been submitted as 'supporting information' due to their size.

Evidence synthesis

A total of 25 PROMs were identified from the 66 publications; 20 disease-specific, 3 generic and 2 utility PROMs (See Table 3). As the included studies were conducted in pre-dialysis, dialysis and renal transplant populations, the evidence synthesis is described in 3 corresponding sections (See Table 5). The term 'indeterminate' was used when vital information required to assess a measurement property was missing (See Table 1) while 'unknown' was used for measurement properties that were only assessed by studies of poor methodological quality (See Table 2).

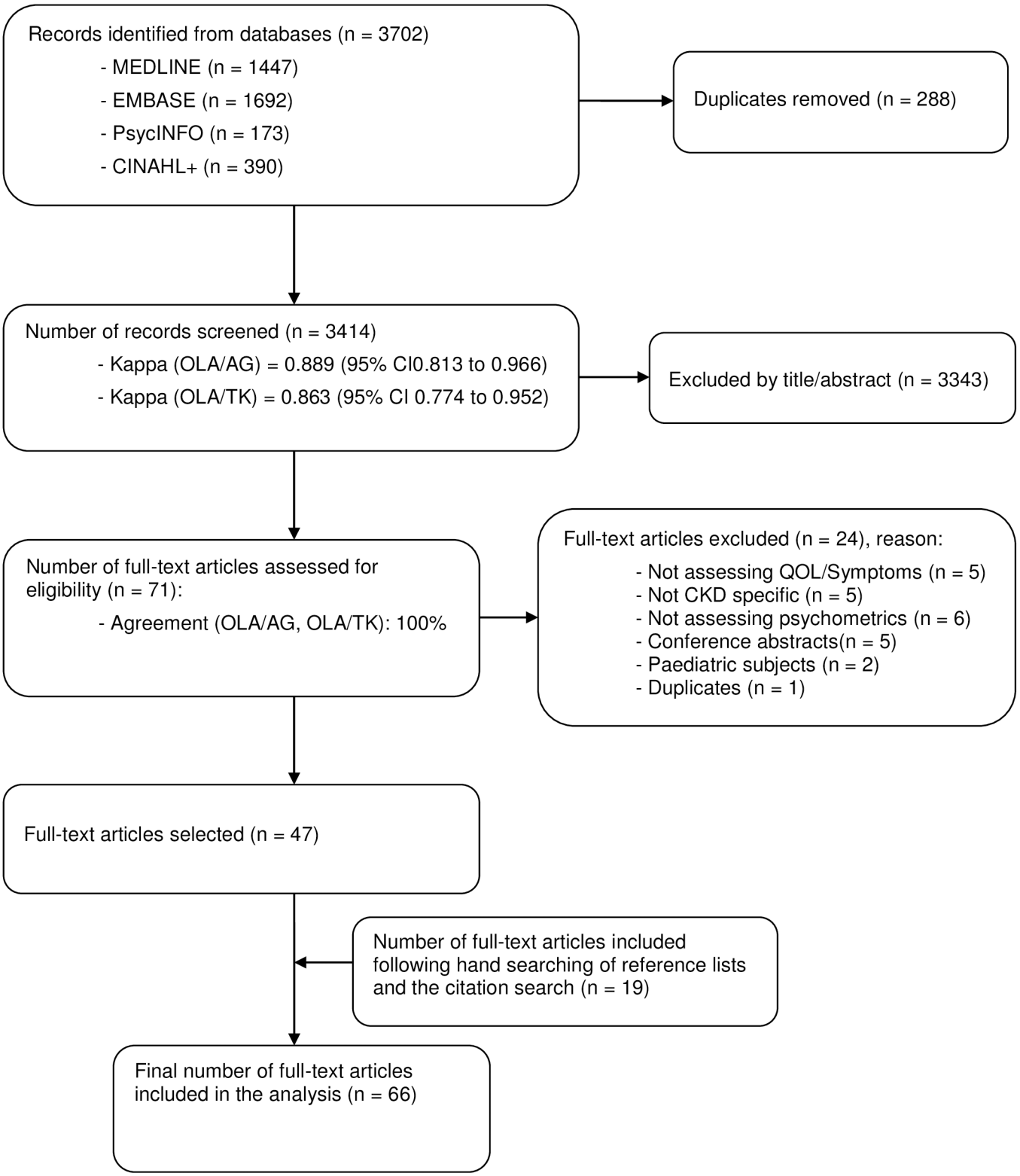

Fig 1. Flow diagram.

https://doi.org/10.1371/journal.pone.0179733.g001

Table 3. Description of PROMs evaluated.

Measure	Description
Disease-specific measures: These measure health in a way that is specific to a particular disease, set of conditions, or part of the body [122].	
Agarwal	A 37-item HRQOL measure for use in non-dialysis patients with mental and physical dimensions [32]. Overall scores range from 0 to 100 with higher scores indicating better HRQOL [32].
Kidney Disease Quality of Life—36 (KDQOL-36)	A 36-item HRQOL measure designed for use in kidney disease patients undergoing dialysis. Derived from the KDQOL-SF [33]. There are 3 specific dimensions namely: symptoms and problems (ii) burden of kidney disease (iii) effects of kidney disease. It also includes two summary scales derived from the SF-12 namely: the physical (PCS) and mental (MCS) scales [34]. Overall scores range from 0 to 100 with higher scores indicating better HRQOL [33].
KDQOL-SF	An 80-item HRQOL measure designed for use in kidney disease patients undergoing dialysis [47]. Derived from the 134-item KDQOL [47, 58]. Version 1.3 differs from version 1.2 by the addition of a screening question for sexual activity [47]. There are 8 generic dimensions from the SF-36 (See below) and 8 disease-specific dimensions namely: (i) symptoms/problems (ii) effects of kidney disease on daily life (iii) burden of kidney disease (iv) work status (v) cognitive function (vi) quality of social interaction (vii) sexual function (viii) sleep. There are 3 additional dimensions namely: (i) social support (ii) dialysis staff encouragement (ii) patient satisfaction. Scores range from 0 to 100 for each dimension and higher scores indicate better HRQOL.
Chinese Dialysis Quality of Life Scale (CDQOL)	A 29-item measure designed to measure the QOL of Chinese dialysis patients. Scored on a 5-point Likert scale. Higher scores indicate better quality of life as perceived by the patient [49].
CHOICE Health Experience Questionnaire (CHEQ)	An 83-item HRQOL measure. Designed to complement the SF-36 and assess the effectiveness of dialysis modalities [50]. Comprises of 8 dimensions from SF36 (see below) combined with 7 supplementary items and 13 specific dimensions namely: (i) freedom (ii) travel restrictions (iii) cognitive functioning (iv) financial (v) restrictions on diet and fluids (vi) recreation (vii) work (viii) body image (ix) symptoms (x) sleep (xi) sexual functioning (xii) access-related problems (xiii) quality of life. An additional 2 person-specific quality of life item. Possible scores range from 0 to 100 with higher scores indicating better HRQOL [50].
Dialysis Symptom Index (DSI)	A 30-item measure designed to assess symptom prevalence and severity in patients on haemodialysis [52]. An overall symptom burden score and a total symptom severity score are calculated. Symptoms are rated on a 5-point Likert scale and total score ranges from 0 to 150 with higher scores indicating greater symptom severity [52].
Modified Edmonton Symptom Assessment System (modified ESAS)	This is a measure of symptom burden for use in dialysis patients. It is a modification of the ESAS. There are 10 symptom-specific items and 10 visual analogue scales with superimposed 0–10 scale [55]. The scale for each symptom is anchored by the words 'No' and 'Severe' at 0 and 10, respectively and the sum of scores range from 0 to 100 with higher scores indicating greater symptom distress and burden [55].

(*Continued*)

Table 3. (*Continued*)

Measure	Description
Kidney Disease Questionnaire (KDQ)	A 26-item measure designed to assess the effects of interventions on the QOL of patients undergoing haemodialysis [56]. There are 5 dimensions namely: (i) physical symptoms (ii) fatigue (iii) depression (iv) relationships with others (v) frustration. Note that the physical symptom dimension is patient-specific, thus the symptoms most important to individual patients are identified and used to evaluate the dimension. Questions are scored on a 7-point Likert scale and higher scores indicate a lower impact of disease on HRQOL [56].
KDQOL	A 134-item QOL measure designed for use in kidney disease patients undergoing dialysis [58]. It consists of SF36 dimensions (see below), 11 kidney disease targeted scales and an item that assesses change in health over a year (overall health rating) [58]. All scale scores are transformed linearly into 0–100 point scales with higher scores indicating better HRQOL [58].
KDQOL (modified)	A 55-item QOL measure derived from the KDQOL [59]. Using affinity mapping, 11 subscales [59] were identified namely: (i) pain (ii) psychological dependency (iii) cognitive functioning (iv) social functioning (v) dialysis-related symptoms (vi) cardiopulmonary symptoms (vii) sleep (viii) energy (ix) cramps (x) diet (xi) appetite. 4 items were ungrouped. The measure is scored on a 0 to 100 scale with higher scores indicating better HRQOL [59].
WHOQOL-BREF (Dialysis)	A 32-item HRQOL measure modified for use in dialysis patients with 4 domains (incorporating 4 dialysis-specific items): (i) physical (ii) psychological (iii) social relationship (iv) environment. The measure also includes two global items (general QOL and general health). A 5-point Likert scale is used and higher scores signify higher HRQOL [63].
Quality of Life Index (QLI) 3.0	A 68-item QOL measure divided into 2 sections. One section measures satisfaction with various domains of life, while the second measures the importance of the domain to the individual [64]. There are 4 domains: (i) health and functioning (ii) Social and economic (iii) psychological/spiritual (iv) family. Each section has 3 additional dialysis-related items. The total QOL score and the four subscale scores range between 0 and 30 with higher scores indicating a better HRQOL [64].
End-Stage Renal Disease Symptom Checklist-Transplantation Module (ESRD-SCL-TM)	A 43-item symptom-specific QOL measure designed for use in renal transplant patients on immunosuppression therapy [70]. There are 6 dimensions: (i) limited physical capacity (ii) limited cognitive capacity (iii) cardiac and renal dysfunction (iv) side effects of corticosteroids (v) increased growth of gum and hair (vi) transplantation-related psychological distress. A 5-point Likert scale is used. Higher scores indicate worse QOL/symptoms [70].
Modified Transplant Symptom Occurrence and Symptom Distress Scale (MTSOSD)	A 29-item measure designed to assess side effects of immunosuppression therapy in renal transplant patients [74]. The symptom occurrence dimension has 20 items while the symptom distress dimension has 9 [74]. Ridit analysis, a statistical method of analysing ordinal data [123] was chosen for data analysis [74]. Higher ridit scores indicate greater symptom distress [74].

(*Continued*)

Table 3. (*Continued*)

Measure	Description
Gastrointestinal Symptom Rating Scale (GSRS)	A 15-item measure that assesses important gastrointestinal side effects of immunosuppressive therapy in renal transplant patients [75]. There are 5 dimensions: (i) reflux (ii) diarrhoea (iii) constipation (iv) abdominal pain (v) indigestion. Each dimension gives an average score ranging from 1 (no discomfort) to 7 (very severe discomfort) with higher scores indicating worse impact [75].
Gastrointestinal Quality of Life Index (GIQLI)	A 36-item measure which focuses on the impact GI complaints on the HRQOL of renal transplant patients [75]. There are 5 dimensions: (i) GI symptoms (ii) emotional status (iii) physical function (iv) social function (v) strain of medical treatment. Total scores range from 0 to 144 with higher scores indicating a better HRQOL [75].
Kidney Transplant Questionnaire (KTQ)	A 25-item HRQOL measure for use in renal transplant patients [78]. It has 5 dimensions namely: (i) physical symptoms (ii) fatigue (iii) uncertainty/fear (iv) appearance (v) emotional. The physical symptom dimension is patient-specific, thus the symptoms most important to individual patients are identified [78]. For each dimension, an average score ranging from 1 to 7 is calculated with higher scores indicating better HRQOL in patients [78].
ReTransQoL (RTQ) version 1	A 45-item measure designed to assess QOL in renal transplant patients [80]. There are 5 dimensions: (i) physical health (ii) mental health (iii) medical care (iv) fear of losing graft (v) treatment. All dimensions are linearly transformed to a 0 to 100 scale and higher scores indicate better HRQOL [80].
ReTransQoL (RTQ) version 2	A 32-item measure designed to assess QOL in renal transplant patients [80]. There are 5 dimensions: (i) physical health (ii) social functioning (iii) medical care (iv) treatment (v) fear of losing graft. All dimensions are linearly transformed to a 0 to 100 scale and higher scores indicate better HRQOL [80].
CKD-Symptom Burden Index (CKD-SBI)	A 32-item measure of symptom burden. Derived from the DSI. The CKD-SBI was developed for use in patients with CKD stages IV and V however it was used in pre-dialysis and dialysis populations in this study [82]. The measure has 4 dimensions namely: (i) prevalence (ii) distress (iii) severity (iv) frequency. Total score ranges from 0 to 100 and higher scores indicate higher symptom burden [82].
Generic measures: These measure health in a general manner and can be used for various health conditions [122].	
Nottingham Health Profile (NHP)	A generic HRQOL questionnaire with 38 yes/no questions [124] grouped into 6 dimensions: (i) pain (ii) energy (iii) physical mobility (iv) sleep (v) emotional reactions (vi) social isolation. There is an optional part. The NHP scores range between 0 (good health status) and 100 (poor health status) [124].
SF-36 version 2	A generic 36-item HRQOL measure with 8 scales [125] namely: (i) physical functioning (ii) physical role (iii) bodily pain (iv) general health (v) vitality (vi) social functioning (vii) emotional role (viii) mental health. An additional 1-item measure of self-evaluated change in health status is available. The Likert rating method is used and raw scores are linearly transformed into 0 to 100 scales with higher transformed scores indicating better HRQOL [125].

(*Continued*)

Table 3. (*Continued*)

Measure	Description
SF-12	A generic 12-item HRQOL measure derived from the SF-36 [126] (see above). The 8 dimensions can be computed into 2 distinct clusters, PCS-12 and MCS-12 with higher values indicating better HRQOL [126].
Utility measures: These provide utilities or values regarding health and can be used for cost-utility analyses of interventions [18].	
EQ-5D	A utility measure with a self-classifier and a visual analogue scale (VAS) which can be used to value health states [127]. The self-classifier includes 5 dimensions: (i) mobility (ii) self-care (iii) usual activities (iv) pain/ discomfort (v) anxiety/depression. Each dimension has 3 levels of severity (no problems, some problems, and severe problems) and it is possible to describe 243 health states between 0 (dead) and 1 (perfect health) [127].
Modified Time Trade-Off (TTO)	This utility measure was used to measure quality of life in dialysis and renal transplant populations and values range from 0 (death) to 1 (full health) [83, 128].

https://doi.org/10.1371/journal.pone.0179733.t003

Pre-dialysis population. Three disease-specific PROMs were used in pre-dialysis populations namely the: Agarwal [32], KDQOL-36 [33, 34] and KDQOL-SF [35, 36]. Although the studies in this section measured estimated glomerular filtration rates (eGFR), there was significant disparity in their description of patients at this stage of CKD. Three studies described patients as 'non-dialysis' [32], 'mild-to-moderate' CKD patients [34] and ESRD patients [36] respectively, while only the studies by Chao et al. [33] and Abd Elhafeez [35] formally categorized patients into CKD stages. None of the studies validated PROMs separately by stage of CKD pre-dialysis.

Agarwal. There was limited evidence for test-retest reliability and content validity. Internal consistency and structural validity were rated as 'unknown' [32].

Kidney Disease Quality of Life—36 (KDQOL-36). Strong evidence for internal consistency was found [33] and there was moderate evidence for hypothesis testing [33, 34]. Structural validity was rated as 'indeterminate' [33].

KDQOL-SF. There was moderate evidence for test-retest reliability [35] and hypothesis testing [35, 36]. Internal consistency, structural validity and responsiveness were rated as 'unknown' [35, 36].

Dialysis population. Fourteen PROMS were used in this group of patients. Among the five studies that evaluated KDQOL-36, two had mixed samples [37, 38]. As the majority of the participants in Tao et al [38] were on dialysis and no significant transplant specific symptoms were elicited in Chow et al. [37] a pragmatic decision was made to analyse them here.

Klersy et al. [39] used a sample which comprised of 85% dialysis and 15% pre-dialysis patients to assess the KDQOL-SF. Again, as the majority were dialysis patients, this study was analysed in this section.

KDQOL-36. Moderate evidence was found for internal consistency [40], test-retest reliability [41, 42] and hypothesis testing [37, 38]. However, structural validity was rated as 'indeterminate' [40].

KDQOL-SF. There was strong evidence of internal consistency and structural validity [43]. Moderate evidence was found for test-retest reliability [44, 45] and hypothesis testing [46] while there was limited evidence for content validity [47]. Responsiveness was rated as 'unknown' [48].

Table 4. Methodological quality of included studies.

Instrument	Study	Language	CKD population	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Translation	Criterion validity	Responsiveness
Agarwal	Agarwal [32]	English	Pre-dialysis	POOR	FAIR		FAIR	POOR	FAIR			
CDQOL	Suet-Ching [49]	Cantonese	Dialysis	POOR	POOR		FAIR		FAIR			
CHEQ	Aiyasanon [51]	Thai	Dialysis	POOR					FAIR	GOOD		
	Wu [50]	English	Dialysis	POOR			GOOD		GOOD			
CKD-SBI	Almutary [82]	Arabic	Pre-dialysis, Dialysis	POOR					FAIR	FAIR		
DSI	Önsoz [53]	Turkish	Dialysis	POOR	FAIR					GOOD		
	Weisbord [52]	English	Dialysis		POOR		FAIR					
QLI 3.0	Dehesh [65]	Persian	Dialysis	POOR				POOR	FAIR	GOOD		
	Ferrans [64]	English	Dialysis	POOR	POOR				FAIR			
	Halabi [66]	Arabic	Dialysis	POOR						FAIR		
	Korkut [67]	Turkish	Dialysis	POOR	FAIR				FAIR	FAIR		
KDQOL (D)	Hays [58]	English	Dialysis	POOR			FAIR	POOR	FAIR			
KDQOL (M)	Rao [59]	English	Dialysis	POOR					FAIR			
KDQOL-36	Chao [33]	Taiwanese	Pre-dialysis	EXCELLENT				EXCELLENT	FAIR	GOOD		
	Chow [37]	Cantonese	Dialysis	POOR	POOR				GOOD			
	Mateti [42]	Kannada	Dialysis	FAIR	FAIR				FAIR	GOOD		
	Ricardo [34]	English, Spanish	Pre-dialysis	POOR					GOOD			
	Tao [38]	Mandarin	Dialysis	FAIR	POOR				FAIR			
	T'charoen [41]	Thai	Dialysis	POOR	FAIR				FAIR			
	Yang [40]	English	Dialysis	GOOD				GOOD	GOOD			
KDQOL-SF	Abd ElHafeez [35]	Arabic	Pre-dialysis	POOR	GOOD			POOR	FAIR	GOOD		
	Barotfi [76]	Hungarian	Transplant	POOR	FAIR				FAIR			
	Bataclan [88]	Filipino	Dialysis	POOR	FAIR				FAIR	GOOD		
	Boini [89]	French	Dialysis	POOR	FAIR				FAIR			
	Bouidida [46]	Moroccan	Dialysis	POOR	POOR				GOOD	GOOD		
	Cheung [36]	Chinese	Pre-dialysis	POOR					FAIR			POOR
	Duarte [96]	Portuguese	Dialysis		FAIR				FAIR	GOOD		
	Fardinmehr [94]	Persian	Dialysis	POOR					FAIR	GOOD		
KDQOL-SF	Green [85]	Japanese	Dialysis	POOR	GOOD				FAIR	GOOD		
	Hays [47]	English	Dialysis	POOR			FAIR					
	Joshi [43]	English, Mandarin, Malay	Dialysis	EXCELLENT				EXCELLENT	GOOD			
	Klersy [39]	Italian	Dialysis	POOR	POOR			POOR	FAIR	GOOD		
	Kontodimopoulos [86]	Greek	Dialysis	POOR					FAIR			
	Kontodimopoulos [87]	Greek	Dialysis	POOR					FAIR			
	Korevaar [48]	Dutch	Dialysis	POOR					GOOD	GOOD		POOR
	Malindretos [44]	Greek	Dialysis	POOR	FAIR				FAIR	GOOD		
	Molsted [92]	Danish	Dialysis	POOR						GOOD		
	Moreira [93]	Portuguese	Dialysis	POOR					FAIR			
	Pakpour [102]	Farsi	Dialysis	FAIR	POOR			FAIR	FAIR	GOOD		
	Park [45]	Korean	Dialysis	POOR	FAIR				POOR	GOOD		
	Perneger [90]	French	Dialysis	POOR					FAIR			
	Vasilieva [95]	Russian	Dialysis	POOR								
	Yildirim [91]	Turkish	Dialysis	POOR					FAIR	FAIR		
KDQ	Alvarez-Ude [57]	Spanish	Dialysis	POOR	FAIR				FAIR	GOOD		

(*Continued*)

Table 4. (*Continued*)

Instrument	Study	Language	CKD population	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Translation	Criterion validity	Responsiveness
	Laupacis [56]	English	Dialysis		FAIR		FAIR	POOR	FAIR			POOR
ESRD-SCL	Franke [70]	German	Transplant	GOOD	POOR		FAIR	GOOD	FAIR			
	Ortega [71]	Spanish	Transplant	GOOD	FAIR			GOOD	FAIR			POOR
	Stavem [72]	Norwegian	Transplant	FAIR	FAIR				GOOD	GOOD		
GIQLI	Kleinman [75]	German, English	Transplant	POOR					GOOD			
GSRS	Kleinman [75]	German, English	Transplant	POOR					GOOD			
KTQ	Chisholm-Burns [84]	English	Transplant	POOR					FAIR			
	Laupacis [78]	English	Transplant	POOR	POOR		FAIR	POOR	FAIR			POOR
	Niu [77]	Chinese	Transplant	GOOD	FAIR			GOOD	FAIR	GOOD		
	Rebollo [79]	Spanish	Transplant	POOR	POOR				FAIR	POOR		POOR
ESAS	Davison [55]	English	Dialysis		FAIR		FAIR		FAIR			
	Davison [54]	English	Dialysis									POOR
MTSOSD	Moons [74]	Dutch	Transplant						GOOD			
RTQ v1	Beauger [81]	French	Transplant					EXCELLENT				
	Gentile [80]	French	Transplant	POOR	POOR		GOOD	POOR	FAIR			POOR
RTQ v2	Beauger [81]	French	Transplant	GOOD				GOOD	FAIR			
EQ-5D	Cleemput [73]	Dutch, French	Transplant						GOOD			
SF-12	Pakpour [62]	Persian	Dialysis	GOOD	GOOD			GOOD	FAIR			
SF-36 v2	Feurer [69]	English	Transplant	POOR					FAIR			
	Mingardi [68]	Italian	Dialysis	GOOD					FAIR			
NHP	Badia [61]	Spanish	Dialysis	FAIR	FAIR							
	Zengin [60]	Turkish	Dialysis	FAIR					FAIR			
TTO (modified)	Churchill [83]	English	Dialysis, Transplant		FAIR				FAIR			
WHOQOL-BREF (D)	Yang [63]	Taiwanese	Dialysis	GOOD	POOR			GOOD	GOOD			

https://doi.org/10.1371/journal.pone.0179733.t004

Table 5. Evidence synthesis of PROMs used in patients with CKD.

Instrument version	Population	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypothesis testing	Criterion validity	Responsiveness
Agarwal [32]	Pre-dialysis	?	+		+	?	-		
KDQOL-36 [33, 34, 37, 38, 40–42]	Pre-dialysis	+++				?*	++		
	Dialysis	++	++			?*	++		
KDQOL-SF [35, 36, 39, 43–48, 76, 85–96, 102]	Pre-dialysis	?	++			?	++		?
	Dialysis	+++	++		+	+++	++		?
	Transplant	?	-				+		
CDQOL [49]	Dialysis	?	?		+		+		
CHEQ [50, 51]	Dialysis	?			++		++		
DSI [52, 53]	Dialysis	?	?		+				
ESAS [54, 55]	Dialysis		+		+		+		?
KDQ [56, 57]	Dialysis	?	+		+	?	++		?
KDQOL (D)[58]	Dialysis	?			+	?			
KDQOL (M)[59]	Dialysis	?					+		
NHP [60, 61]	Dialysis	++	?				+		
SF-12 [62]	Dialysis	++	++			+	+		
WHOQOL-BREF (D) [63]	Dialysis	++	?			?	-		
QLI 3.0 [64–67]	Dialysis	?	-			?	++		
SF-36 v2	Dialysis [68]	++					+		
	Transplant [69]	?					-		
ESRD-SCL [70–72]	Transplant	+++	++		+	++	++		?
EQ-5D [73]	Transplant						++		
GIQLI [75]	Transplant	?					++		
GSRS [75]	Transplant	?					++		
KTQ [77–79, 84]	Transplant	++	+		+	?	++		?
MTSOSD [74]	Transplant						++		
RTQ v1 [80, 81]	Transplant	?	?		++	+/-	+		?
RTQ v2 [81]	Transplant	++				++	+		
TTO (modified) [83]	Mixed (D & TX)		+				+		
CKD-SBI [82]	Mixed (D & Pre-D)	?					+		

+ = positive rating

? = unknown rating

- = negative rating

+/- = conflicting findings

?* = indeterminate rating (due to non-reporting of variance explained by factors)

https://doi.org/10.1371/journal.pone.0179733.t005

Chinese Dialysis Quality of Life Scale (CDQOL). Reliability (internal consistency and test-retest) for CDQOL was rated as 'unknown' and there was limited evidence for content validity and hypothesis testing [49].

CHOICE Health Experience Questionnaire (CHEQ). Moderate evidence for hypothesis testing and content validity was found [50] while internal consistency was rated as 'unknown' [50, 51].

Dialysis Symptom Index (DSI). There was limited evidence for content validity [52]. Reliability (internal consistency and test-retest) was rated as 'unknown' [52, 53].

Modified Edmonton Symptom Assessment System (modified ESAS). There was limited evidence of test-retest reliability, content validity and hypothesis testing. Responsiveness was rated as 'unknown' [54, 55].

Kidney Disease Questionnaire (KDQ). There was moderate evidence of hypothesis testing [56, 57] while there was limited evidence for test-retest reliability and content validity [56]. Structural validity, responsiveness [56] and internal consistency [57] were rated as 'unknown'.

KDQOL. There was limited evidence for content validity while internal consistency and structural validity were rated as 'unknown' [58].

KDQOL (Modified). There was limited evidence for hypothesis testing while internal consistency was rated as 'unknown' [59].

Nottingham Health Profile (NHP). There was moderate evidence for internal consistency [60, 61] and limited evidence for hypothesis testing [60]. Test-retest reliability was rated as 'unknown' [61].

SF-12. Moderate evidence was found for internal consistency and test-retest reliability. There was limited evidence for structural validity and hypothesis testing [62].

WHOQOL-BREF (Dialysis). There was moderate evidence for internal consistency while test-retest reliability and structural validity were rated as 'unknown' [63].

Quality of Life Index (QLI) 3.0. There was moderate evidence for hypothesis testing [64, 65]. Internal consistency and structural validity were rated as 'unknown' [64–67] while there was limited evidence against test-retest reliability [67].

SF-36 version 2. Moderate evidence was found for internal consistency while limited evidence was found for hypothesis testing [68].

Renal transplant population. Ten PROMs were evaluated specifically in renal transplant patients and all except the SF-36 and the EQ-5D were disease specific measures.

SF-36 version 2. Internal consistency was rated as 'unknown' [69].

End-Stage Renal Disease Symptom Checklist–Transplantation Module (ESRD-SCL-TM). There was strong evidence for internal consistency [70, 71], moderate evidence for test-retest reliability [71, 72], hypothesis testing [72] and structural validity [71]. There was limited evidence for content validity [70] while responsiveness was rated as 'unknown' [71].

EQ-5D. Moderate evidence for hypothesis testing was found [73].

Modified Transplant Symptom Occurrence and Symptom Distress Scale (MTSOSD). Moderate evidence for hypothesis testing was found for MTSOSD [74].

Gastrointestinal Symptom Rating Scale (GSRS). Moderate evidence for hypothesis testing was found while internal consistency was rated as 'unknown' [75].

Gastrointestinal Quality of Life Index (GIQLI). Moderate evidence for hypothesis testing was found while internal consistency was rated as 'unknown' [75].

KDQOL-SF. There was limited evidence for hypothesis testing and limited evidence against test-retest reliability. Internal consistency was rated as 'unknown' [76].

Kidney Transplant Questionnaire (KTQ). There was moderate evidence for internal consistency [77] and hypothesis testing [78, 79]. Limited evidence was found for test-retest reliability [77] and content validity [78] Structural validity [77] and responsiveness [79] were rated as 'unknown'.

ReTransQoL (RTQ) version 1. There was moderate evidence of content validity [80], conflicting evidence of structural validity [80, 81] and limited evidence of hypothesis testing [80]. Internal consistency, test-retest reliability and responsiveness were rated as 'unknown' [80].

RTQ version 2. This revised version had moderate evidence for internal consistency and structural validity. There was still limited evidence for hypothesis testing [81].

Single studies with mixed samples. The study by Almutary et al. [82] evaluated the CKD-Symptom Burden Index (CKD-SBI) in a mixed sample of pre-dialysis and dialysis patients while Churchill et al. [83] assessed the Time Trade-Off (TTO) in a mixed sample of dialysis and renal transplant recipients.

CKD-Symptom Burden Index (CKD-SBI)

Limited evidence was found for hypothesis testing while internal consistency was rated as 'unknown' [82].

Modified Time Trade-Off (TTO)

There was limited evidence for test-retest reliability and hypothesis testing [83].

Other findings

Reliability. Internal consistency: Of the 58 studies that assessed internal consistency, 42 were scored 'poor' for methodological quality. This was due to one of 3 reasons:

1. Some studies did not conduct a factor analysis and did not reference a relevant study that conducted one [41, 49–51, 53, 61, 64, 75, 82].
2. Some studies conducted a factor analysis but had inadequate sample sizes by COSMIN standards [32, 35, 39, 47, 59, 65, 78, 80].
3. Some studies referenced a relevant study that conducted a factor analysis but the sample size used for the study was inadequate by COSMIN standards [34, 36, 37, 44–48, 57, 66, 67, 69, 76, 79, 84–95].

Test-retest reliability: Thirty-four studies conducted test-retest reliability and most of them reported internal correlation coefficients (ICC). The study by Duarte et al. [96] was the only one that reported inter & intra-observer reliability. The majority scored 'fair' for test-retest reliability and this was largely due to the small sample sizes.

Measurement error: The included studies did not provide adequate information on parameters such as the minimal important change (MIC) [97], the standard error of measurement (SEM) [30, 98] or the limits of agreement (LOA) [99] making it difficult to assess measurement error. Only one study provided an estimate for minimal clinical important difference (MCID) [75].

Validity. Content validity: This was assessed by development studies for RTQ [80], modified ESAS [55], Agarwal [32], CHEQ [50], CDQOL [49], DSI [52], KDQ [56], KTQ [78], KDQOL [58], KDQOL-SF [47]. Six validation studies reported content validity indexes (CVI) [33, 38, 49, 53, 77, 82] and of these only Suet-Ching [49] reported patient involvement in the process of content validation.

Construct validity (Hypothesis testing): Of the 66 studies evaluated, only 13 reported clearly formulated a priori hypotheses or expectations regarding the magnitude and direction of correlations. For this reason, most of the studies were rated as 'fair' for the methodological quality of their hypothesis testing.

The absence of clear a priori hypotheses make it difficult to determine whether any results reported for construct validity and responsiveness was due to chance or not [100, 101].

Structural validity: A number of measures were rated poorly for structural validity due to issues with factor analysis (See internal consistency). Out of the 19 studies that conducted a factor analysis, 8 were scored 'poor' [32, 35, 39, 56, 58, 65, 78, 80] and this can be attributed to

the use of sample sizes inadequate by COSMIN standards (n < 5 times the total number of items and < 100) [100]. It is important that studies perform factor analysis as it verifies scale structure and uni-dimensionality which determines the scoring and interpretation of a measure's internal consistency statistic [100].

Criterion validity: This was not assessed for any study as the COSMIN Delphi panel does not regard any PROM as true 'gold standard' [29] and the FDA holds a similar view [14].

Cross-cultural validity/Translations: Twenty-five studies translated PROM instruments and adapted them to varying degrees for their study population. As none of these studies performed a multi-group confirmatory factor analysis or assessed differential item functioning (DIF) between language groups, the decision was made not to assess cross-cultural validity. Therefore, only the quality of their translations was assessed according to the provisions specified in the COSMIN manual [100].

The translations conducted by 20 studies [33, 35, 39, 42, 44–46, 48, 51, 53, 57, 65, 72, 77, 85, 88, 92, 94, 96, 102] were rated as good, while the translations by 4 studies [66, 67, 82, 91] were rated as fair and 1 translation was rated as poor [79].

Responsiveness. Responsiveness was only assessed by 8 of the included studies [36, 48, 54, 56, 71, 78–80] and all were rated 'poor' as the information provided was inadequate.

Interpretability. A number of the PROMs included in this review had significant floor and ceiling effects which might indicate a reduction in their ability to discriminate between patients with the lowest or highest possible scores and detect changes over time [30].

Ten instruments (KDQOL-SF, Agarwal, KDQ, KDQOL-36, KTQ, QLI, KDQOL, ESRD-SCL, CHEQ and NHP) had floor and ceiling effects > 15% while the RTQ v2 and SF-12 [62, 81] had < 15%.

The 'work status' dimension of the KDQOL-SF had a 70% floor effect [45, 86] while the 'social support', 'patient satisfaction' and 'staff encouragement' domains had ceiling effects of 46%, 45% and 67% [86], respectively. Sexual function had ceiling effects of 53.3% while cognitive function had ceiling effects of 60% [35]. The 'pain' dimension of the SF-36 had ceiling effects as high as 59% [35].

Eight studies namely; RTQ v1 [80], RTQv2 [81], GSRS & GIQLI [75], KDQOL-36 [38, 41], KDQOL-SF [44], CHEQ [51], TTO [83] reported measurement scores for subgroups within their study populations. While some of the differences in scores were statistically significant, it is unclear if any were clinically relevant.

Feasibility and acceptability. This was difficult to assess for the included studies as less than a third reported the average time needed to complete the questionnaires, few reported the recall period used and none reported administrative requirements for collection and analysis of data. However, most of the studies reported good response rates which suggest that patients might find the use of PROMs acceptable.

The fact that majority of the studies failed to report the level and/or the method used for handling missing data meant there might be a risk of bias [101]. Questions relating to sexual activities had the highest levels of missing data with Bataclan and Dial [88] reporting a response rate of <18% in Filipino patients [88].

Patient involvement. Patient involvement in the process of item generation and/or item selection was reported in the development of the RTQ [80], KDQ [56], KTQ [78], ESRD-SCLTM [70], KDQOL-SF [47], Agarwal [32], CHEQ [50], QLI 3.0 [64], KDQOL (Dialysis) [58] and CDQOL [49].

All the studies that translated and adapted measures reported pre-testing their translations in patients to assess a combination of comprehension, cultural relevance and acceptability except Rebollo et al.[79]

Discussion

This is the first review to use the COSMIN checklist [100] to evaluate the measurement properties of PROMs used in patients with CKD. In all, 25 PROMs were evaluated by a total of 66 studies in pre-dialysis, dialysis, and renal transplant patients.

In the pre-dialysis population, the KDQOL-36 exhibited strong evidence for internal consistency and moderate evidence for construct validity (hypothesis testing). It should be noted that the evidence for this measure was obtained from studies conducted in Taiwanese patients [33], and a combination of Hispanic and non-Hispanic white patients [34]. Furthermore, the measurement properties were not reported by CKD stage. Therefore, further validation may be necessary before use outside these study populations and/or where focus is on the use of PROMs in relationship to the severity of CKD.

In dialysis patients, we found evidence to support the use of both the KDQOL-SF and the KDQOL-36. The KDQOL-SF demonstrated strong evidence for internal consistency and structural validity and moderate evidence for test-retest reliability and construct validity (hypothesis testing) while the KDQOL-36 had moderate evidence for internal consistency, test-retest reliability and construct validity (hypothesis testing). Again it should be noted that this evidence was obtained from a significant number of non-English studies, further validation work would be needed before these measures could be confidently utilised in an English speaking population.

In renal transplant patients, the ESRD-SCLTM demonstrated strong evidence for internal consistency and moderate evidence for test-retest reliability, structural validity and construct validity (hypothesis testing).

Consistent with the review by Gibbons and Fitzpatrick [18], much of the evidence we present was derived from cross-sectional studies. However, in contrast to that study, we did not include clinical trial reports as the methodological quality of their PROM evaluations are often unsatisfactory [23] and their PROM analysis often inadequate and insufficiently [103] reported for any meaningful evaluation to be possible [19].

In line with our findings, Gibbons and Fitzpatrick [18] found evidence to support the use of the KDQOL-SF but did not specify which modality of renal replacement therapy (RRT) provided the evidence. We found evidence to support the use of the KDQOL-36 (which was not available at the time Gibbons and Fitzpatrick [18] conducted their review), and the ESRD-SCLTM (which was excluded from their study) [18]. The ERA-EDTA expert panel [104] recommended the KDQOL-36, following a consensus meeting.

There were methodological issues with the majority of the PROMs included in this review. These included: sample sizes smaller than current recommendations, a lack of clearly described a priori hypotheses and inadequate reporting of missing data, and little or no information on measurement error and responsiveness. Similar issues were reported by Gibbons and Fitzpatrick [18] and highlighted in reviews for other health conditions [105–107].

Given that the COSMIN standards only became available within the last decade, it is unsurprising that most of the earlier studies fared poorly when judged against these exacting methodological and reporting standards even though they might meet the minimum standards recommended by the International Society for Quality of Life Research (ISOQOL) [108]. However, this highlights the need to test and revise PROM instruments on a regular basis to ensure they actually perform as intended according to contemporary psychometric standards.

This systematic review provides a basis for identifying PROMs with potential utility in clinical practice. There is evidence that the use of PROMs in clinical practice could enhance communication [109–112] between patients and their clinicians. Basch et al. [113] noted that the use of PROMs in routine clinical care, could facilitate the reporting of serious adverse events due to drug toxicities [113], while a review by Finkelstein et al. [114] suggested that the use of

PROMs could assist renal teams with the development of strategies to improve the HRQOL of the patient with CKD [114]. Calvert et al. [115] suggested that PROM data could potentially facilitate the delivery of tailored healthcare if successfully integrated with routinely collected clinical and laboratory data [115].

Whilst effective management of risk factors can slow CKD progression [116] many patients with severe pre-dialysis CKD progress to end stage renal disease (ESRD) and in this group PROMs may have a significant role [117–119]. For example, PROMs could be used to monitor individual patients for symptoms or changes in HRQOL that may indicate that a medical review or intervention is needed [17, 120].

Whilst a systematic review can provide valuable evidence on psychometric properties of PROMs, clinicians and researchers need to consider a number of issues such as the domains covered by different measures, the available supporting evidence, and the suitability for the target population and use (whether for clinical trials, routine practice, audit or real-time decision making). As no single measure covers all the domains that might be of interest, there might be a need to administer more than one measure. Patient acceptability is also a key issue. Therefore, it is important that patients are involved not just in the development of measures but also in the selection for research and/or practice.

During the course of this review, we became aware of the existence of the IPOS-Renal and contacted its developers [121]. This measure is currently being used within the measurement work stream of the UK Renal Registry and its validation by the Palliative care Outcome Scale (POS) team is on-going [121]. Therefore, evidence to support its use may be available in future.

The key strengths of this review are the use of the COSMIN standards and criteria for evidence synthesis which ensured that our assessments of the included PROMs were robust, the stratification of the review by stage of CKD and the absence of language restrictions which strengthened our findings. The main limitation is the fact that the included studies did not adequately report their assessments for a number of measurement properties thus making it difficult and sometimes impossible to evaluate these properties.

At present, we suggest the use of the KDQOL-36 in pre-dialysis patients though initial validation might be required. We recommend the KDQOL-SF and the KDQOL-36 for use in dialysis patients as we found evidence supporting both of these measures. The shorter 36-item KDQOL-36 may be more practical for use in routine clinic settings, while the longer 80-item KDQOL-SF might be preferred for research purposes where more detailed information may be required. We suggest using the ESRD-SCLTM in renal transplant recipients to assess issues pertaining to renal transplantation and immunosuppression therapy. These measures are recommended based on the fact that they currently possess the best evidence available according to COSMIN standards in these populations and meet the minimum standards recommended by ISOQOL [108]. However, it must be recognised that none of these measures possess evidence underpinning *all* measurement properties and some lack validation in English-speaking populations. Future work should be undertaken to address these gaps. For all measures, it is vital that content validity is established according to FDA guidelines to ensure that they actually measure the concept(s) of interest. This should be conducted before other measurement properties are fully evaluated and adequately reported in order to facilitate their subsequent evaluation. Investigators may use this review to identify the gaps in evidence and design studies to address these issues in future.

Supporting information

S1 Text. Review protocol.
(PDF)

S2 Text. PRISMA checklist.
(DOC)

S3 Text. Search strategy.
(PDF)

S1 Table. Summary of study results.
(DOCX)

S2 Table. Characteristics of included studies.
(DOCX)

S3 Table. Characteristics of included PROMs.
(DOCX)

Acknowledgments

We are grateful to Susan Bayliss who assisted with the development of the search strategy [24]; Dr Lazaros Andronis, and Dr Fouad Al-Baaj who translated some papers; and Dr Harjeet Bhachu who assisted with the tables.

Author Contributions

Conceptualization: MC DK PC TM OLA.

Data curation: OLA TK AG AS.

Formal analysis: OLA AS.

Funding acquisition: MC DK PC TM.

Investigation: OLA TK AG AS.

Methodology: MC DK PC TM OLA.

Project administration: MC DK PC TM OLA.

Supervision: MC DK PC TM.

Visualization: MC DK PC TM OLA.

Writing – original draft: OLA.

Writing – review & editing: MC DK PC TM OLA TK AG AS.

References

1. Couser WG, Remuzzi G, Mendis S, Tonelli M. The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. Kidney international. 2011; 80(12):1258–70. Epub 2011/10/14. https://doi.org/10.1038/ki.2011.368 PMID: 21993585.
2. Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, de Jong PE, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet (London, England). 2010; 375 (9731):2073–81. Epub 2010/05/21. https://doi.org/10.1016/s0140-6736(10)60674-5 PMID: 20483451; PubMed Central PMCID: PMCPmc3993088.
3. Nitsch D, Grams M, Sang Y, Black C, Cirillo M, Djurdjev O, et al. Associations of estimated glomerular filtration rate and albuminuria with mortality and renal failure by sex: a meta-analysis. BMJ (Clinical research ed). 2013; 346:f324. Epub 2013/01/31. https://doi.org/10.1136/bmj.f324 PMID: 23360717; PubMed Central PMCID: PMCPmc3558410.

4. Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al. Global Prevalence of Chronic Kidney Disease—A Systematic Review and Meta-Analysis. PloS one. 2016; 11(7):e0158765. Epub 2016/07/08. https://doi.org/10.1371/journal.pone.0158765 PMID: 27383068; PubMed Central PMCID: PMCPmc4934905.
5. Hamer RA, El Nahas AM. The burden of chronic kidney disease: Is rising rapidly worldwide. BMJ: British Medical Journal. 2006; 332(7541):563–4. PMC1397782. https://doi.org/10.1136/bmj.332.7541.563 PMID: 16528062
6. Kerr M, Bray B, Medcalf J, O'Donoghue DJ, Matthews B. Estimating the financial cost of chronic kidney disease to the NHS in England. Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association—European Renal Association. 2012; 27 Suppl 3:iii73–80. Epub 2012/07/21. https://doi.org/10.1093/ndt/gfs269 PMID: 22815543; PubMed Central PMCID: PMCPmc3484716.
7. Honeycutt AA, Segel JE, Zhuo X, Hoerger TJ, Imai K, Williams D. Medical Costs of CKD in the Medicare Population. Journal of the American Society of Nephrology. 2013; 24(9):1478–83. https://doi.org/10.1681/ASN.2012040392 PMID: 23907508
8. Meyer T, Hostetter T. Medical Progress: Uremia. The New England journal of medicine. 2007; 13 (357):1316–25.
9. Jablonski A. The multidimensional characteristics of symptoms reported by patients on hemodialysis. Nephrol Nurs J. 2007; 34(1):29–37; quiz 8. PMID: 17345690.
10. Almutary H, Bonner A, Douglas C. Symptom burden in chronic kidney disease: a review of recent literature. Journal of Renal Care. 2013; 39(3):140–50. http://dx.doi.org/10.1111/j.1755-6686.2013.12022.x. PMID: 23826803. https://doi.org/10.1111/j.1755-6686.2013.12022.x
11. Merkus MP, Jager KJ, Dekker FW, de Haan RJ, Boeschoten EW, Krediet RT. Physical symptoms and quality of life in patients on chronic dialysis: results of The Netherlands Cooperative Study on Adequacy of Dialysis (NECOSAD). Nephrology Dialysis Transplantation. 1999; 14(5):1163–70. https://doi.org/10.1093/ndt/14.5.1163
12. Calvert MJ, Freemantle N. Use of health-related quality of life in prescribing research. Part 1: why evaluate health-related quality of life? Journal of clinical pharmacy and therapeutics. 2003; 28(6):513–21. Epub 2003/12/04. PMID: 14651676.
13. Kyte DG, Draper H, Ives J, Liles C, Gheorghe A, Calvert M. Patient reported outcomes (PROs) in clinical trials: is 'in-trial' guidance lacking? a systematic review. PloS one. 2013; 8(4):e60684. Epub 2013/04/06. https://doi.org/10.1371/journal.pone.0060684 PMID: 23560103; PubMed Central PMCID: PMCPmc3613381.
14. FDA. Patient-reported outcome measures: use in medicinal product development to support labeling claims. Guidance for industry. 2009. Available: https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf. Accessed 22 May 2017.
15. Black N. Patient reported outcome measures could help transform healthcare. BMJ (Clinical research ed). 2013; 346:f167. Epub 2013/01/30. https://doi.org/10.1136/bmj.f167 PMID: 23358487.
16. Chris Ham CI, Nick Goodwin, Anna Dixon, Patrick South. Where next for the NHS reforms? The case for integrated care. The King's Fund. 2011.
17. Hjollund NH, Larsen LP, Biering K, Johnsen SP, Riiskjaer E, Schougaard LM. Use of Patient-Reported Outcome (PRO) Measures at Group and Patient Levels: Experiences From the Generic Integrated PRO System, WestChronic. Interactive journal of medical research. 2014; 3(1):e5. Epub 2014/02/13. https://doi.org/10.2196/ijmr.2885 PMID: 24518281; PubMed Central PMCID: PMCPmc3936283.
18. Gibbons E, Fitzpatrick, R. A structured review of patient-reported outcome measures for people with chronic kidney disease. Department of Public Health University of Oxford. 2010.
19. Terwee CB. Protocol for systematic reviews of measurement properties. COSMIN. 2011.
20. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2012; 21(4):651–7. Epub 2011/07/07. https://doi.org/10.1007/s11136-011-9960-1 PMID: 21732199; PubMed Central PMCID: PMCPmc3323819.
21. Moher D, Tsertsvadze A, Tricco AC, Eccles M, Grimshaw J, Sampson M, et al. When and how to update systematic reviews. The Cochrane database of systematic reviews. 2008;(1):Mr000023. Epub 2008/02/07. https://doi.org/10.1002/14651858.MR000023.pub3 PMID: 18254126.
22. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2010; 19(4):539–49. Epub 2010/02/20.

https://doi.org/10.1007/s11136-010-9606-8 PMID: 20169472; PubMed Central PMCID: PMCPmc2852520.

23. Martini C, Gamper EM, Wintner L, Nilica B, Sperner-Unterweger B, Holzner B, et al. Systematic review reveals lack of quality in reporting health-related quality of life in patients with gastroenteropancreatic neuroendocrine tumours. Health and quality of life outcomes. 2016; 14(1):127. Epub 2016/09/12. https://doi.org/10.1186/s12955-016-0527-2 PMID: 27614762; PubMed Central PMCID: PMCPmc5018190.

24. Aiyegbusi OL, Kyte D, Cockwell P, Marshall T, Keeley T, Gheorghe A, et al. Measurement properties of patient-reported outcome measures (PROMs) used in adult patients with chronic kidney disease: a systematic review protocol. BMJ Open. 2016; 6(10):e012014. Epub 2016/10/14. https://doi.org/10.1136/bmjopen-2016-012014 PMID: 27733411.

25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ (Clinical research ed). 2009; 339:b2700. Epub 2009/07/23. https://doi.org/10.1136/bmj.b2700 PMID: 19622552; PubMed Central PMCID: PMCPmc2714672.

26. Terwee CB, Prinsen CA, Ricci Garotti MG, Suman A, de Vet HC, Mokkink LB. The quality of systematic reviews of health-related outcome measurement instruments. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2016; 25(4):767–79. Epub 2015/09/09. https://doi.org/10.1007/s11136-015-1122-4 PMID: 26346986; PubMed Central PMCID: PMCPmc4830864.

27. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2009; 18(8):1115–23. Epub 2009/08/28. https://doi.org/10.1007/s11136-009-9528-5 PMID: 19711195; PubMed Central PMCID: PMCPmc2744791.

28. PROM Group AM, Carolina Casañas i Comabella, Monica Hadi, Elizabeth Gibbons, Ray Fitzpatrick, Nia Roberts. PROM GROUP CONSTRUCT & INSTRUMENT TYPE FILTERS. 2010.

29. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. Journal of clinical epidemiology. 2010; 63(7):737–45. Epub 2010/05/25. https://doi.org/10.1016/j.jclinepi.2010.02.006 PMID: 20494804.

30. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of clinical epidemiology. 2007; 60(1):34–42. Epub 2006/12/13. https://doi.org/10.1016/j.jclinepi.2006.03.012 PMID: 17161752.

31. Cohen J. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement. 1960; 20(1):37–46. https://doi.org/10.1177/001316446002000104

32. Agarwal R. Developing a self-administered CKD symptom assessment instrument. Nephrology Dialysis Transplantation. 2010; 25(1):160–6. http://dx.doi.org/10.1093/ndt/gfp426. PMID: 2010005312.

33. Chao S, Yen M, Lin TC, Sung JM, Wang MC, Hung SY. Psychometric Properties of the Kidney Disease Quality of Life-36 Questionnaire (KDQOL-36). West J Nurs Res. 2016; 38(8):1067–82. Epub 2016/03/31. https://doi.org/10.1177/0193945916640765 PMID: 27026384.

34. Ricardo AC, Hacker E, Lora CM, Ackerson L, De Salvo KB, Go A, et al. Validation of the Kidney Disease Quality of Life Short Form 36 (KDQOL-36TM) US Spanish and English versions in a cohort of Hispanics with chronic kidney disease. Ethnicity and Disease. 2013; 23(2):202–9. PMID: 23530302.

35. Abd Elhafeez S, Sallam SA, Gad ZM, Zoccali C, Torino C, Tripepi G, et al. Cultural adaptation and validation of the "kidney Disease and Quality of Life—Short Form (KDQOL-SFTM) version 1.3" questionnaire in Egypt. BMC nephrology. 2012; 13 (1) (no pagination)(170). http://dx.doi.org/10.1186/1471-2369-13-170. PMID: 2013024551.

36. Cheung YB, Seow YY, Qu LM, Yee ACP. Measurement properties of the Chinese version of the Kidney Disease Quality of Life-Short Form (KDQOL-SFTM) in end-stage renal disease patients with poor prognosis in singapore. Journal of Pain and Symptom Management. 2012; 44(6):923–32. http://dx.doi.org/10.1016/j.jpainsymman.2011.12.282. PMID: 22795902. https://doi.org/10.1016/j.jpainsymman.2011.12.282

37. Chow SK, Tam BM. Is the kidney disease quality of life-36 (KDQOL-36) a valid instrument for Chinese dialysis patients? BMC nephrology. 2014; 15:199. http://dx.doi.org/10.1186/1471-2369-15-199. PMID: 25511462; PubMed Central PMCID: PMCPMC4274701. https://doi.org/10.1186/1471-2369-15-199

38. Tao X, Chow SK, Wong FK. Determining the validity and reliability of the Chinese version of the Kidney Disease Quality of Life Questionnaire (KDQOL-36). BMC nephrology. 2014; 15:115. Epub 2014/07/13. https://doi.org/10.1186/1471-2369-15-115 PMID: 25015224; PubMed Central PMCID: PMCPmc4115482.

39. Klersy C, Callegari A, Giorgi I, Sepe V, Efficace E, Politi P. Italian translation, cultural adaptation and validation of KDQOL-SF, version 1.3, in patients with severe renal failure. J Nephrol. 2007; 20(1):43–51. Epub 2007/03/10. PMID: 17347972.

40. Yang F, Wang VW, Joshi VD, Lau TW, Luo N. Validation of the English version of the Kidney Disease Quality of Life questionnaire (KDQOL-36) in haemodialysis patients in Singapore. Patient. 2013; 6 (2):135–41. http://dx.doi.org/10.1007/s40271-013-0015-2. PMID: 23613285. https://doi.org/10.1007/s40271-013-0015-2

41. Thaweethamcharoen T, Srimongkol W, Noparatayaporn P, Jariyayothin P, Sukthinthai N, Aiyasanon N, et al. Validity and reliability of KDQOL-36 in thai kidney disease patient. Value in Health Regional Issues. 2013; 2(1):98–102. http://dx.doi.org/10.1016/j.vhri.2013.02.011. PMID: 2013289571.

42. Mateti UV, Nagappa AN, Attur RP, Nagaraju SP, Mayya SS, Balkrishnan R. Cross-cultural adaptation, validation and reliability of the South Indian (Kannada) version of the Kidney Disease and Quality of Life (KDQOL-36) instrument. Saudi J Kidney Dis Transpl. 2015; 26(6):1246–52. Epub 2015/11/21. https://doi.org/10.4103/1319-2442.168662 PMID: 26586066.

43. Joshi VD, Mooppil N, Lim JFY. Validation of the kidney disease quality of life-short form: A cross-sectional study of a dialysis-targeted health measure in singapore. BMC nephrology. 2010; 11 (1) (no pagination)(36). http://dx.doi.org/10.1186/1471-2369-11-36. PMID: 2011020886.

44. Malindretos P, Sarafidis P, Spaia S, Sioulis A, Zeggos N, Raptis V, et al. Adaptation and validation of the Kidney Disease Quality of Life-Short Form questionnaire in the Greek language. Am J Nephrol. 2010; 31(1):9–14. Epub 2009/10/30. https://doi.org/10.1159/000252926 PMID: 19864884.

45. Park HJ, Kim S, Yong JS, Han SS, Yang DH, Meguro M, et al. Reliability and validity of the Korean version of kidney disease quality of life instrument (KDQOL-SFTM). Tohoku Journal of Experimental Medicine. 2007; 211(4):321–9. http://dx.doi.org/10.1620/tjem.211.321. PMID: 17409671.

46. Rhou H, Ezaitouni F, Ouzeddoun N, Bayahia R, Elhajji K, Roudies R, et al. Translation, cultural adaptation and validation of the kidney disease quality of life-short form 1.3 in an African country. Transplantation Proceedings. 2014; 46(5):1295–301. http://dx.doi.org/10.1016/j.transproceed.2014.02.011. PMID: 24935292. https://doi.org/10.1016/j.transproceed.2014.02.011

47. Hays R. D. KJD, Mapes D. L., Coons S. J., Amin N., Carter W. B., Kamberg C. Kidney Disease Quality of Life Short Form (KDQOL-SF) Version 1.3: A manual for use and scoring. RAND. 1995.

48. Korevaar JC, Merkus MP, Jansen MAM, Dekker FW, Boeschoten EW, Krediet RT. Validation of the KDQOL-SFTM: A dialysis-targeted health measure. Quality of Life Research. 2002; 11(5):437–47. http://dx.doi.org/10.1023/A:1015631411960. PMID: 12113391.

49. Suet-Ching WL. The psychometric properties of the Chinese Dialysis Quality of Life Scale for Hong Kong dialysis patients. Journal of advanced nursing. 2001; 36(3):441–9. PMID: 11686759.

50. Wu AW, Fink NE, Cagney KA, Bass EB, Rubin HR, Meyer KB, et al. Developing a health-related quality-of-life measure for end-stage renal disease: The CHOICE Health Experience Questionnaire. American journal of kidney diseases: the official journal of the National Kidney Foundation. 2001; 37(1):11–21. Epub 2001/01/03. PMID: 11136162.

51. Aiyasanon N, Premasathian N, Nimmannit A, Jetanavanich P, Sritippayawan S. Validity and reliability of CHOICE health experience questionnaire: Thai version. Journal of the Medical Association of Thailand. 2009; 92(9):1159–66. PMID: 19772174.

52. Weisbord SD, Fried LF, Arnold RM, Rotondi AJ, Fine MJ, Levenson DJ, et al. Development of a symptom assessment instrument for chronic hemodialysis patients: The Dialysis Symptom Index. Journal of Pain and Symptom Management. 2004; 27(3):226–40. http://dx.doi.org/10.1016/j.jpainsymman.2003.07.004. PMID: 15010101. https://doi.org/10.1016/j.jpainsymman.2003.07.004

53. Onsoz H. UYO. Reliability and validity of the Turkish version of the Dialysis Symptom Index in chronic hemodialysis patients. Turkish Nephrology, Dialysis and Tranplantation Journal. 2013; 22(1):60–7.

54. Davison SN, Jhangri GS, Johnson JA. Longitudinal validation of a modified Edmonton symptom assessment system (ESAS) in haemodialysis patients. Nephrology Dialysis Transplantation. 2006; 21 (11):3189–95. http://dx.doi.org/10.1093/ndt/gfl380. PMID: 2006504906.

55. Davison SN, Jhangri GS, Johnson JA. Cross-sectional validity of a modified Edmonton symptom assessment system in dialysis patients: a simple assessment of symptom burden. Kidney international. 2006; 69(9):1621–5. PMID: 16672923. https://doi.org/10.1038/sj.ki.5000184

56. Laupacis A, Muirhead N, Keown P, Wong C. A disease-specific questionnaire for assessing quality of life in patients on hemodialysis. Nephron. 1992; 60(3):302–6. Epub 1992/01/01. PMID: 1565182.

57. Alvarez-Ude F, Galan P, Vicente E, Alamo C, Fernandez-Reyes MJ, Badia X. Adaptation and validation of the Spanish version of 'Kidney Disease Questionnaire'. [Spanish]. Nefrologia. 1997; 17(6):486–96. PMID: 1998037857.

58. Hays RD, Kallich JD, Mapes DL, Coons SJ, Carter WB. Development of the kidney disease quality of life (KDQOL) instrument. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 1994; 3(5):329–38. Epub 1994/10/01. PMID: 7841967.

59. Rao S, Carter WB, Mapes DL, Kallich JD, Kamberg CJ, Spritzer KL, et al. Development of subscales from the symptoms/problems and effects of kidney disease scales of the Kidney Disease Quality of Life Instrument. Clinical Therapeutics. 2000; 22(9):1099–111. http://dx.doi.org/10.1016/S0149-2918%2800%2980087-9. PMID: 11048907. https://doi.org/10.1016/S0149-2918(00)80087-9

60. Zengin N, Oren B, Gul A, Ustundag H. Assessment of quality of life in haemodialysis patients: A comparison of the Nottingham Health Profile and the Short Form 36. International Journal of Nursing Practice. 2014; 20(2):115–25. http://dx.doi.org/10.1111/ijn.12130. PMID: 24713007. https://doi.org/10.1111/ijn.12130

61. Badia X, Alonso J, Brosa M, Lock P. Reliability of the Spanish version of the Nottingham health profile in patients with stable end-stage renal disease. Social Science and Medicine. 1994; 38(1):153–8. http://dx.doi.org/10.1016/0277-9536%2894%2990310-7. PMID: 8146706.

62. Pakpour AH, Nourozi S, Molsted S, Harrison AP, Nourozi K, Fridlund B. Validity and reliability of short form-12 questionnaire in Iranian hemodialysis patients. Iran J Kidney Dis. 2011; 5(3):175–81. Epub 2011/04/29. PMID: 21525577.

63. Yang SC, Kuo PW, Wang JD, Lin MI, Su S. Development and pyschometric properties of the dialysis module of the WHOQOL-BREF Taiwan version. Journal of the Formosan Medical Association. 2006; 105(4):299–309. PMID: 16618610.

64. Ferrans CE, Powers MJ. Quality of life index: development and psychometric properties. Ans. 1985; Advances in nursing science. 8(1):15–24. 3933411.

65. Dehesh T, Zare N, Jafari P, Sagheb MM. Psychometric assessment of the Persian version of the Ferrans and Powers 3.0 index in hemodialysis patients. International Urology and Nephrology. 2014; 46 (6):1183–9. http://dx.doi.org/10.1007/s11255-013-0537-5. PMID: 23979815. https://doi.org/10.1007/s11255-013-0537-5

66. Halabi JO. Psychometric properties of the Arabic version of Quality of Life Index. Journal of Advanced Nursing. 2006; 55(5):604–10. http://dx.doi.org/10.1111/j.1365-2648.2006.03952_1.x. PMID: 16907792. https://doi.org/10.1111/j.1365-2648.2006.03952_1.x

67. Korkut Y. The reliability and validity study of the Turkish version of Ferrans and Powers' Quality of Life Index for dialysis patients. Archives of Neuropsychiatry. 2007; 44(14).

68. Mingardi G, Cornalba L, Cortinovis E, Ruggiata R, Mosconi P, Apolone G. Health-related quality of life in dialysis patients. A report from an Italian study using the SF-36 Health Survey. Nephrology Dialysis Transplantation. 1999; 14(6):1503–10. http://dx.doi.org/10.1093/ndt/14.6.1503. PMID: 1999183110.

69. Feurer ID, Moore DE, Speroff T, Liu H, Payne J, Harrison C, et al. Refining a health-related quality of life assessment strategy for solid organ transplant patients. Clinical Transplantation, Supplement. 2004; 18(12):39–45. PMID: 2004392633.

70. Franke GH, Reimer J, Kohnle M, Luetkes P, Maehner N, Heemann U. Quality of life in end-stage renal disease patients after successful kidney transplantation: development of the ESRD symptom checklist—transplantation module. Nephron. 1999; 83(1):31–9. Epub 1999/08/26. 45470. PMID: 10461033. https://doi.org/45470

71. Ortega T, Valdes C, Rebollo P, Ortega F. Evaluation of reliability and validity of Spanish version of the end-stage renal disease symptom checklist-transplantation module. Transplantation. 2007; 84 (11):1428–35. http://dx.doi.org/10.1097/01.tp.0000290231.39240.df. PMID: 18091518. https://doi.org/10.1097/01.tp.0000290231.39240.df

72. Stavem K, Ganss R. Reliability and validity of the ESRD Symptom Checklist—Transplantation Module in Norwegian kidney transplant recipients. BMC nephrology. 2006; 7:17. Epub 2006/11/18. https://doi.org/10.1186/1471-2369-7-17 PMID: 17109746; PubMed Central PMCID: PMCPmc1660533.

73. Cleemput I, Kesteloot K, Moons P, Vanrenterghem Y, Van Hooff JP, Squifflet JP, et al. The construct and concurrent validity of the EQ-5D in a renal transplant population. Value in Health. 2004; 7(4):499–509. https://doi.org/10.1111/j.1524-4733.2004.74013.x PMID: 15449642.

74. Moons P, De Geest S, Versteven K, Abraham I, Vlaminck H, Moens G, et al. Psychometric properties of the "Modified Transplant Symptom Occurrence and Symptom Distress Scale". Journal of nursing measurement. 2001; 9(2):115–34. PMID: 11696937.

75. Kleinman L, Kilburg A, Machnicki G, Faull R, Walker R, Prasad R, et al. Using GI-specific patient outcome measures in renal transplant patients: Validation of the GSRS and GIQLI. Quality of Life Research. 2006; 15(7):1223–32. http://dx.doi.org/10.1007/s11136-006-0053-5. PMID: 16972162. https://doi.org/10.1007/s11136-006-0053-5

76. Barotfi S, Molnar MZ, Almasi C, Kovacs AZ, Remport A, Szeifert L, et al. Validation of the Kidney Disease Quality of Life-Short Form questionnaire in kidney transplant patients. Journal of Psychosomatic

Research. 2006; 60(5):495–504. http://dx.doi.org/10.1016/j.jpsychores.2005.09.009. PMID: 16650590. https://doi.org/10.1016/j.jpsychores.2005.09.009

77. Niu Y, Zhang W, Mao S, Gao Y, Wang J, Li J, et al. Pilot feasibility research of Chinese version of kidney transplant questionnaire in recipients of living donor kidney transplantation. Int J Clin Exp Med. 2015; 8(12):22570–6. Epub 2016/02/18. PMID: 26885244; PubMed Central PMCID: PMCPmc4730030.

78. Laupacis A, Pus N, Muirhead N, Wong C, Ferguson B, Keown P. Disease-specific questionnaire for patients with a renal transplant. Nephron. 1993; 64(2):226–31. Epub 1993/01/01. PMID: 8321355.

79. Rebollo P, Ortega F, Ortega T, Valdes C, Garcia-Mendoza M, Gomez E. Spanish validation of the "kidney transplant questionnaire": a useful instrument for assessing health related quality of life in kidney transplant patients. Health and quality of life outcomes. 2003; 1:56. Epub 2003/11/14. https://doi.org/10.1186/1477-7525-1-56 PMID: 14613566; PubMed Central PMCID: PMCPmc269994.

80. Gentile S, Jouve E, Dussol B, Moal V, Berland Y, Sambuc R. Development and validation of a French patient-based health-related quality of life instrument in kidney transplant: The ReTransQoL. Health and quality of life outcomes. 2008; 6 (no pagination)(78). http://dx.doi.org/10.1186/1477-7525-6-78. PMID: 2008534167.

81. Beauger D, Gentile S, Jouve E, Dussol B, Jacquelinet C, Briancon S. Analysis, evaluation and adaptation of the ReTransQoL: A specific quality of life questionnaire for renal transplant recipients. Health and quality of life outcomes. 2013; 11 (1) (no pagination)(148). http://dx.doi.org/10.1186/1477-7525-11-148. PMID: 2013572155.

82. Almutary H, Bonner A, Douglas C. Arabic translation, adaptation and modification of the dialysis symptom index for chronic kidney disease stages four and five. BMC nephrology. 2015; 16(1):1–8. https://doi.org/10.1186/s12882-015-0036-2 PMID: 25884303

83. Churchill DN, Torrance GW, Taylor DW, Barnes CC, Ludwin D, Shimizu A, et al. Measurement of quality of life in end-stage renal disease: the time trade-off approach. Clin Invest Med. 1987; 10(1):14–20. Epub 1987/01/01. PMID: 3545580.

84. Chisholm-Burns MA, Erickson SR, Spivey CA, Gruessner RWG, Kaplan B. Concurrent validity of kidney transplant questionnaire in US renal transplant recipients. Patient Preference and Adherence. 2011; 5:517–22. http://dx.doi.org/10.2147/PPA.S24261. https://doi.org/10.2147/PPA.S24261 PMID: 22114465

85. Green J, Fukuhara S, Shinzato T, Miura Y, Wada S, Hays RD, et al. Translation, cultural adaptation, and initial reliability and multitrait testing of the Kidney Disease Quality of Life Instrument for use in Japan. Quality of Life Research. 2001; 10(1):93–100. http://dx.doi.org/10.1023/A:1016630825992. PMID: 11508479.

86. Kontodimopoulos N, Niakas D. Determining the basic psychometric properties of the Greek KDQOL-SFTM. Quality of Life Research. 2005; 14(8):1967–75. http://dx.doi.org/10.1007/s11136-005-3868-6. PMID: 16155785. https://doi.org/10.1007/s11136-005-3868-6

87. Kontodimopoulos N, Niakas D. Measuring health-related quality of life of Greek dialysis patients with the KDQOL-SFTM. [Greek]. Archives of Hellenic Medicine. 2007; 24(6):590–9. PMID: 2008072477.

88. Bataclan RP, Dial MAD. Cultural adaptation and validation of the Filipino version of Kidney Disease Quality of Life—Short Form (KDQOL-SF version 1.3). Nephrology. 2009; 14(7):663–8. http://dx.doi.org/10.1111/j.1440-1797.2009.01130.x. PMID: 19796025. https://doi.org/10.1111/j.1440-1797.2009.01130.x

89. Boini S, Leplege A, Loos Ayav C, Francais P, Ecosse E, Briancon S. Measuring quality of life in end-stage renal disease. Transcultural adaptation and validation of the specific Kidney Disease Quality of Life questionnaire. [French]. Nephrologie et Therapeutique. 2007; 3(6):372–83. http://dx.doi.org/10.1016/j.nephro.2007.05.005. PMID: 17919640. https://doi.org/10.1016/j.nephro.2007.05.005

90. Perneger TV, Leski M, Chopard-Stoermann C, Martin PY. Assessment of health status in chronic hemodialysis patients. Journal of Nephrology. 2003; 16(2):252–9. PMID: 12768073.

91. Yildirim A, Ogutmen B, Bektas G, Isci E, Mete M, Tolgay HI. Translation, cultural adaptation, initial reliability, and validation of the Kidney Disease and Quality of Life-Short Form (KDQOL-SF 1.3) in Turkey. Transplant Proc. 2007; 39(1):51–4. Epub 2007/02/06. 10.1016/j.transproceed.2006.10.196. PMID: 17275473. https://doi.org/10.1016/j.transproceed.2006.10.196

92. Molsted S, Heaf J, Prescott L, Eidemak I. Reliability testing of the Danish version of the Kidney Disease Quality of Life Short Form. Scand J Urol Nephrol. 2005; 39(6):498–502. Epub 2005/11/24. https://doi.org/10.1080/00365590500240253 PMID: 16303727.

93. Moreira CA, Garletti Junior W, Lima LF, Lima CR, Ribeiro JF, Miranda AF. Assessment of the basic psychometric properties for the portuguese version of the KDQOL-SFTM. [Portuguese]. Rev Assoc Med Bras. 2009; 55(1):22–8. PMID: 19360273.

94. Fardinmehr O, Farajzadegan Z, Naini AE, Mortazavi M, Gholamrezaei A. The validity and reliability of the persian version of kidney disease quality of life questionnaire-short form (KDQOL-SF) in Iranian patients. [Persian]. Journal of Isfahan Medical School. 2012; 29(165). PMID: 2014599849.

95. Vasilieva IA. Russian version of the questionnaire The Kidney Disease and Quality of Life Short Form (KDQOL-SF) a valuable diagnostic instrument for assessing quality of life of dialysis patients. Nefrologiya. 2007; 11(1):64–70.

96. Duarte PS, Ciconelli RM, Sesso R. Cultural adaptation and validation of the "Kidney Disease and Quality of Life—Short Form (KDQOL-SFTM 1.3)" in Brazil. Brazilian Journal of Medical and Biological Research. 2005; 38(2):261–70. PMID: 15785838. https://doi.org//S0100-879X2005000200015

97. Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury. 2011; 42(3):236–40. Epub 2010/12/15. https://doi.org/10.1016/j.injury.2010.11.042 PMID: 21145544.

98. Terwee CB, Roorda LD, Knol DL, De Boer MR, De Vet HC. Linking measurement error to minimal important change of patient-reported outcomes. Journal of clinical epidemiology. 2009; 62(10):1062–7. Epub 2009/02/24. https://doi.org/10.1016/j.jclinepi.2008.10.011 PMID: 19230609.

99. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England). 1986; 1(8476):307–10. Epub 1986/02/08. PMID: 2868172.

100. Mokkink LB, Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L. et al. The COSMIN checklist manual. 2009.

101. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol. 2010; 10:22. Epub 2010/03/20. https://doi.org/10.1186/1471-2288-10-22 PMID: 20298572; PubMed Central PMCID: PMCPmc2848183.

102. Pakpour AH, Yekaninejad M, Molsted S, Harrison AP, Hashemi F, Saffari M. Translation, cultural adaptation assessment, and both validity and reliability testing of the Kidney Disease Quality of Life—Short Form version 1.3 for use with Iranian patients. Nephrology. 2011; 16(1):106–12. http://dx.doi.org/10.1111/j.1440-1797.2010.01389.x. PMID: 21175986. https://doi.org/10.1111/j.1440-1797.2010.01389.x

103. Bjordal K. Impact of quality of life measurement in daily clinical practice. Annals of oncology: official journal of the European Society for Medical Oncology. 2004; 15 Suppl 4:iv279–82. Epub 2004/10/13. 10.1093/annonc/mdh939. PMID: 15477321. https://doi.org/10.1093/annonc/mdh939

104. Breckenridge K, Bekker HL, Gibbons E, van der Veer SN, Abbott D, Briançon S, et al. How to routinely collect data on patient-reported outcome and experience measures in renal registries in Europe: an expert consensus meeting. Nephrology Dialysis Transplantation. 2015. https://doi.org/10.1093/ndt/gfv209 PMID: 25982327

105. Elbers RG, Rietberg MB, van Wegen EE, Verhoef J, Kramer SF, Terwee CB, et al. Self-report fatigue questionnaires in multiple sclerosis, Parkinson's disease and stroke: a systematic review of measurement properties. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2012; 21(6):925–44. Epub 2011/10/21. https://doi.org/10.1007/s11136-011-0009-2 PMID: 22012025; PubMed Central PMCID: PMCPmc3389599.

106. Green A, Liles C, Rushton A, Kyte DG. Measurement properties of patient-reported outcome measures (PROMS) in Patellofemoral Pain Syndrome: a systematic review. Man Ther. 2014; 19(6):517–26. Epub 2014/07/07. https://doi.org/10.1016/j.math.2014.05.013 PMID: 24997774.

107. Janssens A, Rogers M, Gumm R, Jenkinson C, Tennant A, Logan S, et al. Measurement properties of multidimensional patient-reported outcome measures in neurodisability: a systematic review of evaluation studies. Dev Med Child Neurol. 2016; 58(5):437–51. Epub 2015/12/15. https://doi.org/10.1111/dmcn.12982 PMID: 26661333; PubMed Central PMCID: PMCPmc5031226.

108. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2013; 22(8):1889–905. Epub 2013/01/05. https://doi.org/10.1007/s11136-012-0344-y PMID: 23288613.

109. Gilbody SM, Whitty PM, Grimshaw JM, Thomas RE. Improving the detection and management of depression in primary care. Quality & safety in health care. 2003; 12(2):149–55. Epub 2003/04/08. PMID: 12679514; PubMed Central PMCID: PMCPmc1743696. https://doi.org/10.1136/qhc.12.2.149

110. Velikova G, Booth L, Smith AB, Brown PM, Lynch P, Brown JM, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. Journal of clinical oncology: official journal of the American Society of Clinical Oncology. 2004; 22(4):714–24. Epub 2004/02/18. https://doi.org/10.1200/jco.2004.06.078 PMID: 14966096.

111. Mathias SD, Fifer SK, Mazonson PD, Lubeck DP, Buesching DP, Patrick DL. Necessary but not sufficient: the effect of screening and feedback on outcomes of primary care patients with untreated anxiety. J Gen Intern Med. 1994; 9(11):606–15. Epub 1994/11/01. PMID: 7853069.

112. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. Journal of evaluation in clinical practice. 2006; 12(5):559–68. Epub 2006/09/22. https://doi.org/10.1111/j.1365-2753.2006.00650.x PMID: 16987118.

113. Basch E, Deal AM, Kris MG, Scher HI, Hudis CA, Sabbatini P, et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. Journal of Clinical Oncology. 2016; 34(6):557–65. https://doi.org/10.1200/JCO.2015.63.0830 PMID: 26644527.

114. Finkelstein FO, Wuerth D, Finkelstein SH. Health related quality of life and the CKD patient: challenges for the nephrology community. Kidney international. 76(9):946–52. https://doi.org/10.1038/ki.2009.307 PMID: 19675529

115. Calvert M, Thwaites R, Kyte D, Devlin N. Putting patient-reported outcomes on the 'Big Data Road Map'. Journal of the Royal Society of Medicine. 2015; 108(8):299–303. Epub 2015/04/02. https://doi.org/10.1177/0141076815579896 PMID: 25827908.

116. Program NKDE. Making sense of CKD—A Consise Guide for Managing Chronic Kidney Disease in the Primary Care Setting. NIH Publication. 2014;(14–7989).

117. Murtagh FEM, Sheerin NS, Addington-Hall J, Higginson IJ. Trajectories of Illness in Stage 5 Chronic Kidney Disease: A Longitudinal Study of Patient Symptoms and Concerns in the Last Year of Life. Clinical Journal of the American Society of Nephrology. 2011; 6(7):1580–90. https://doi.org/10.2215/CJN.09021010 PMID: 21685021

118. Tangri N, Stevens LA, Griffith J, Tighiouart H, Djurdjev O, Naimark D, et al. A predictive model for progression of chronic kidney disease to kidney failure. Jama. 2011; 305(15):1553–9. Epub 2011/04/13. https://doi.org/10.1001/jama.2011.451 PMID: 21482743.

119. Yu HT. Progression of chronic renal failure. Archives of internal medicine. 2003; 163(12):1417–29. Epub 2003/06/26. https://doi.org/10.1001/archinte.163.12.1417 PMID: 12824091.

120. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2009; 18(1):115–23. Epub 2008/12/24. https://doi.org/10.1007/s11136-008-9430-6 PMID: 19105048.

121. POS. Palliative care Outcome Scale Website. Available: https://pos-pal.org/ Accessed: 22 May 2017

122. Devlin N, Appleby J. Getting the most out of PROMs: Putting health outcomes at the heart of NHS decision-making. London: Kings Fund and Office of Health Economics. 2010.

123. Irwin DJB. How to Use Ridit Analysis. Biometrics. 1958; 14(1):18–38. https://doi.org/10.2307/2527727

124. Hunt SM, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. The Journal of the Royal College of General Practitioners. 1985; 35(273):185–8. PMID: 3989783.

125. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Medical care. 1992; 30(6):473–83. Epub 1992/06/11. PMID: 1593914.

126. Ware J Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Medical care. 1996; 34(3):220–33. Epub 1996/03/01. PMID: 8628042.

127. Johnson JA, Coons SJ, Ergo A, Szava-Kovats G. Valuation of EuroQOL (EQ-5D) health states in an adult US sample. Pharmacoeconomics. 1998; 13(4):421–33. Epub 1998/03/08. PMID: 10178666.

128. Torrance GW. Measurement of health state utilities for economic appraisal. Journal of health economics. 1986; 5(1):1–30. Epub 1986/02/09. PMID: 10311607.